SciLifeLab

Instructions: *Use Times New Roman font size 11, 1.0 pt text space, and* ***maximum three A4 pages****, including any figures.*

Name or type of instrument	**Sequencing Instrument Next Generation PacBio**
Unit	NGI-UGC
Platform	Genomics
Applicant (Head of Unit)	Susanne Hellstedt Kerje
E-mail address	susanne.kerje@igp.uu.se
Instrument, acquisition cost (kSEK)	8500
Funding applied for (kSEK)	8500

Description of the instrument

Describe the instrument, its purpose, and its specific capabilities. If possible, provide a link to a web page describing the type of instrument. Does the instrument represent a new technology at the unit, or is it a replacement/upgrade of current equipment base? Is the instrument unique in terms of service to users nationally?

The next generation sequencing instrument from PacBio is a single molecule real-time (SMRT) long-read sequencing instrument, expected to be released early 2023. This instrument represents a significant leap forward compared to earlier PacBio instrumentation, with near 10 times higher throughput and with the possibility to run flow cells in parallel, a feature not possible earlier. This system is uniquely flexible as each flow cell is controlled individually giving the instrument the potential to be run continuously with no start-up or shut-down times.

One new feature with this system is the increased sequencing capacity per SMRT cell with 15 times higher throughput per cell and reduced consumable requirements. The new reagents for this system are also automation ready and there are protocols available for running on Hamilton or Agilent liquid handling systems. With this instrument the cost for running a human genome is approaching 1000 USD and can be generated in one day.

Sequencing data generated on the instrument is handled in an integrated GPU based compute resource significantly reducing post-sequencing data handling. The resulting HiFi reads are rapidly processed and the output is a .bam-file, which is an optimal starting point for further bioinformatic analyses.

As the instrument is not available for purchase until spring 2023 there are no description or specifications yet publicly available. The description above is based on information given Uppsala Genome Center under NDA.

As we are the only provider of long-read sequencing using the PacBio technology in Sweden, this new instrument will not be available for service elsewhere.

Statement of need

Describe how the instrument will be crucial for the operations, productivity, development, and competitiveness of your unit. Estimate the need and impact of the instrument from a user perspective. Will the instrument be shared with or used by other units or research groups at SciLifeLab? Also, describe any co-funding plans.

Over the last few years long-read sequencing has reached a level of accuracy and cost that makes it an attractive option for many applications. With the investments in the latest PacBio Sequel II instruments (expensive instrument funding 2020), providing HiFi reads (i.e. long-read sequence data of the same accuracy as Illumina short-read sequencing) we have seen an increased interest and user base in the scientific community. For de novo genome assembly in biodiversity projects, PacBio HiFi data is now the golden standard and is used for all species sequenced in international consortia such as the Earth BioGenome Project/European Reference Genome Atlas (ERGA) and Darwin Tree of Life (1). Meanwhile, the cost of generating PacBio sequence data with the current instrument is still prohibitively expensive for analysis of human genomes at scale and we currently have a limited user base from the medical community.

Here we apply for a new PacBio sequencing system that will be launched in 2023. With the new instrument the cost of generating HiFi data will drop significantly, making it an attractive alternative to standard Illumina sequencing for human genome sequencing projects. The system will be more flexible than Sequel II and will produce significantly more data from the same input material. We therefore see an enormous potential with the new system for the medical community, in addition to lower cost and higher throughput for the biodiversity projects that are already favoring PacBio sequencing.

In addition to whole genome sequencing for human and biodiversity projects a decreased cost of accurate long-read sequencing data will also make PacBio sequencing an attractive application for entirely new user groups. We see single cell long-read sequencing as a new and growing application. The ability to sequence full-length transcripts and isoforms provides added value for understanding the complexity of single cell transcriptomes. The new MAS-Iso-seq protocol, a technique for concatenating cDNAs into single molecules optimal for long-read sequencing, allows boosting the throughput >15 fold compared with standard Iso-seq protocol. With significantly increased throughput and reduced cost of sequencing on the new system we apply for here, we see enormous potential in single cell analysis going forward. Efforts to unify single cell analysis within SciLifeLab are ongoing, and this instrument will thus be a great resource for the larger community.

Investment in the next generation long-read sequencing system from PacBio will be crucial for Uppsala Genome Center and SciLifeLab to retain its world leading position in the long-read sequencing arena. With the launch of the new system, it is our assessment that the current Sequel II instruments will quickly become obsolete for many applications, and our users will demand sequencing on the new system. Depreciation costs for the current Sequel II systems will end by 2023 and it will thus be timely to have a new system up and running during the next year. We expect the new system to rapidly replace current PacBio systems at larger genome cores internationally. Since PacBio sequencing is not typically offered as a commercial product, we see a great need for the Swedish research community to have this system at SciLifeLab.

We currently do not have any co-funding for purchase of the new system. Considering the cost of the system we would therefore require a significant fraction of the instrument cost to be funded in this call to go ahead with the acquisition. We are also willing to look into parallel funding opportunities to cover costs not funded in the expensive instrument call.

Reference

1. Lawniczak et al, Standards recommendations for the Earth BioGenome Project, https://www.pnas.org/doi/10.1073/pnas.2115639118